# Measurement Error Does Not Explain the Persistence of a Body Mass Index Association with Endometrial Cancer after Adjustment for Endogenous Hormones

*Nancy Potischman, Mitchell H. Gail, Rebecca Troisi, Sholom Wacholder, Robert N. Hoover*

Identified risk factors for endometrial cancer are accepted as operating through estrogen exposure. In a recent analysis, the effect of risk factors such as body mass index (BMI) was not explained by circulating estrogen concentrations. In the present analysis, we correct for measurement error associated with obtaining only one blood sample per subject. Applying regression calibration ideas, we found that error correction of log estrone had little impact on estimates of the BMI effect, suggesting that hormone measurement error does not account for the residual importance of BMI. The biologic mechanism for the increased risk associated with BMI remains to be explained. (Epidemiology 1999;10:76–79)

**Keywords: regression calibration, errors in variables, measurement error correction, hormones, endometrial cancer.**

We recently published results on risk related to endogenous hormone levels in a case-control study of endometrial cancer,[1] testing the hypothesis that epidemiologic risk factors for endometrial cancer operate through endogenous hormonal mechanisms, namely circulating estrogens. Support for this hypothesis comes from studies of unopposed estrogen therapy.[2,3] Another risk factor, obesity, is also thought to operate through estrogen mechanisms,[4] since estrogens are produced in adipose tissue.[5,6] These estrogens are also more bioavailable because of lower levels of sex-hormone binding globulin (SHBG) in obese compared with non-obese women.[7,8] Other risk factors, including history of oral contraceptive use, parity, and age at menopause, also are thought to operate through circulating hormones.[9,10] To our knowledge, our study is the first large-scale epidemiologic endeavor to collect both questionnaire and serologic data, and to evaluate the relations among risk factors.

## Subjects and Methods

Our case-control study[1] involved newly diagnosed cases from seven hospitals and controls matched on age, race, and area of residence. Community controls under age 65 were obtained through random-digit dialing procedures; those age 65 or older were derived from the Health Care Financing Administration. Women referred to these hospitals for benign gynecological conditions formed an additional control group. Results were not different for the two control groups, which we combined for these analyses. Subjects were interviewed and then measured for a variety of anthropometric indices including height and weight, from which we calculated the body mass index (BMI kg/m$^2$).[11] We excluded women who reported use of exogenous estrogen within 6 months, resulting in 208 cases and 209 controls for analysis. Assays were conducted by Nichols Institute (San Juan Capistrano, CA).

## Results

Presuming that endogenous estrogen levels explain the effect of risk factors such as BMI in our study, we expected that the odds ratios (ORs) for endogenous estrogens would be of substantially greater magnitude than those for the risk factors. We also anticipated that effect estimates for "surrogate" risk factors would be eliminated or substantially diminished after adjustment for an exposure variable that mediated the effect. Neither of these expectations was met in these data. For example, adjusted for other factors besides hormones, the OR for highest to lowest BMI category was 3.8 (95% confidence interval (CI) = 2.2–6.4), whereas the estimates for the highest *vs* lowest estrogen quartiles ranged from 3.0 for total estradiol to 3.8 for estrone without adjustment for BMI. With adjustment for BMI, the effect estimates were reduced to 1.3 and 2.2 for estradiol and estrone, respectively. In contrast, the BMI odds ratio was essentially unchanged by adjustment for any of the hormone measures. The odds ratio for BMI was barely affected even when adjusted for SHBG ($OR_{BMI}$ = 3.1,

From the Epidemiology and Biostatistics Program, Division of Cancer Epidemiology and Genetics, National Cancer Institute, Bethesda, MD.

Address correspondence to: Nancy Potischman, Nutritional Epidemiology Branch, Division of Cancer Epidemiology and Genetics, National Cancer Institute, Executive Plaza North, Suite 430, Bethesda, MD 20892-7366.

Submitted January 13, 1998; final version accepted May 18, 1998.

95% CI = 1.8–5.4), despite strong associations of SHBG with both BMI and endometrial cancer risk in these data. The effect estimates for other established risk factors (for example parity, smoking) were similarly unaffected after adjustment for endogenous hormones.

Because the hormone levels used in these adjustments were based on a single sample from each woman, we assessed whether errors in estimates of estrogen levels could account for their minimal impact on the risk estimate for BMI. We studied estrone because it is the predominant circulating estrogen in postmenopausal women. To examine the effects of errors in estrone measurement, we considered the logistic model

$$\log\{p/(1-p)\} = \mu + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3, \qquad (1)$$

where $p$ is the risk of endometrial cancer, $X_1$ is the natural logarithm of true recent (assume within 3 years) estrone level (pg/mL) for a given woman, $X_2$ is her BMI, and $X_3$ is her age in years. Both $X_2$ and $X_3$ are assumed to be measured without error. Because we have only one value of estrone, which fluctuates over time for a given woman, we do not measure $X_1$ but rather $Z_1 = X_1 + \epsilon$, where $\epsilon$ is assumed independent of $X_1$. The quantities $X_1$ and $\epsilon$ are assumed normally distributed with respective means $\mu$ and 0 and variances $\sigma_{11}$ and $\sigma^2$. The estimated coefficients, which we obtain by including $Z_1$ instead of the true value $X_1$ in Eq 1 are $\hat{\beta}_1^* = 0.44$, $\hat{\beta}_2^* = 0.081$ and $\hat{\beta}_3^* = 0.025$. Applying regression calibration ideas,[12] we obtain corrected estimates that account for the measurement error in $Z_1$ (Appendix). These corrected estimates are $\hat{\beta}_1 = 0.62$, $\hat{\beta}_2 = 0.075$, $\hat{\beta}_3 = 0.026$. We see that correction for error in $Z_1$ "deattenuates" the odds ratio associated with ln (estrone) considerably, but has only a minor impact in decreasing the apparent effect of BMI from $\hat{\beta}_2^* = 0.081$ to $\hat{\beta}_2 = 0.075$. We conclude that measurement error in estrone does not account for the residual importance of BMI after adjustment for ln (estrone).

## Discussion

There are several alternative explanations for the persistence of BMI as an important factor after adjustment for estrone. First, long-term postmenopausal estrone levels may be the causal factor, and current BMI may better reflect long-term estrone exposure than do current postmenopausal estrone levels. Nevertheless, the observation that excess endometrial cancer risk from exogenous estrogen diminishes within 2–5 years of cessation[13] suggests that recent exposure is important. A second possibility is that BMI is a marker for other factors and that both estrone and these factors contribute independently to risk. One alternative pathway for a BMI effect was thought to be through insulin resistance,[14] which should be a marker for a variety of metabolic aberrations.[15] In these data, however, there was no main effect of c-peptide, a marker of insulin secretion, after adjustment for BMI. In addition, the odds ratio for BMI was unchanged after adjustment for c-peptide.

**TABLE 1. Crude Odds Ratios by Tertile of Estrone (E1) and Quartile of BMI**

| | Estrone (pg/mL) | | |
|---|---|---|---|
| BMI (kg/m²) | <26 | 26–40 | ≥41 |
| <23.0 | 1.00* (15) | 1.49 (12) | 0.78 (5) |
| 23.0–26.0 | 0.87 (8) | 1.07 (12) | 0.93 (6) |
| 26.1–30.0 | 0.73 (7) | 0.70 (9) | 1.87 (11) |
| >30.0 | 0.27 (1) | 6.43 (31) | 4.80 (90) |

Numbers in parentheses indicate number of cases.
* Reference group.

Another possibility is that factors represented by BMI not only act as independent risk factors, as in Eq 1, but also interact with estrogens to elevate risk. We calculated crude odds ratios for tertiles of estrone and quartiles of BMI (Table 1). The data were suggestive of an interaction.

Two assumptions deserve mention. The estimate of intraclass correlation was based on annual measurements over the most recent 2–3 years.[16] It is possible, although we believe unlikely, that errors about a woman's long-term mean are positively correlated over yearly intervals. In this case, a lower intraclass correlation might be found for measurements taken at longer time intervals, such as decades. Nevertheless, even if the intraclass correlation is 0.66, which is the lower confidence limit in the data of Hankinson *et al*,[16] the value of the BMI coefficient is only reduced to 0.072 from 0.081 (an 11% reduction) (Appendix). For the BMI coefficient to be reduced by 50%, the intraclass correlation would have to be 0.3–0.4, which seems implausible. Second, we assumed a classical error model wherein the errors in the estrone measurement were independent of the true estrone value.[17] Perhaps this assumption does not hold precisely, but we are unaware of data to suggest such a violation of the classical model.

Our findings, together with our earlier finding of increased risk associated with high androstenedione concentrations, suggest a need to investigate alternative explanations for the risk associated with BMI and for the etiology of endometrial cancer.

## References

1. Potischman N, Hoover RN, Brinton LA, Siiteri P, Dorgan JF, Swanson CA, Berman ML, Mortel R, Twiggs LB, Barrett RJ, Wilbanks GD, Persky V, Lurain JR: Case-control study of endogenous steroid hormones and endometrial cancer. J Natl Cancer Inst 1996;88:1127–1135.
2. Voigt LF, Weiss NS, Chu J, Daling JR, McKnight B, Van Bell G: Progestagen supplementation of exogenous oestrogens and risk of endometrial cancer. Lancet 1991;338:274–277.
3. Brinton LA, Hoover RN, and the Endometrial Cancer Collaborative Group: Estrogen replacement therapy and endometrial cancer risk: unresolved issues. Obstet Gynecol 1993;81:265–271.
4. Parazzini F, La Vecchia C, Bocciolone L, Franceschi S. The epidemiology of endometrial cancer. Gynecol Oncol 1991;41:1–16.
5. Siiteri PK, Nisker JA, Hammond GL: Hormonal basis of risk factors for breast and endometrial cancer. In: Iacobelli S, King RJB, Lindner HR, Lippmann ME (eds), Hormones and Cancer. New York: Raven Press, 1980; 499–505.

6. MacDonald PC, Edman CD, Hemsell DL, Porter JC, Siiteri PK: Effect of obesity on conversion of plasma androstenedione to estrone in postmenopausal women with and without endometrial cancer. Am J Obstet Gynecol 1978;130:448–455.
7. Nisker JA, Hammond GL, Davidson BJ, Frumar AM, Takami NK, Judd HL, Siiteri PK: Serum sex hormone-binding globulin capacity and the percentage of free estradiol in postmenopausal women with and without endometrial carcinoma. Am J Obstet Gynecol 1980;138:637–642.
8. Nyholm HCJ, Nielsen AL, Lyndrup J, Dreisler A, Hagen C, Haug E: Plasma oestrogens in postmenopausal women with endometrial cancer. Br J Obstet Gynecol 1993;100:1115–1119.
9. Henderson BE, Ross RK, Pike MC, Casagrande JT: Endogenous hormones as major factor in human cancer. Cancer Res 1982;42:3232–3239.
10. Key TJ, Pike MC: The dose-effect relationship between "unopposed" estrogens and endometrial mitotic rate: its central role in explaining and predicting endometrial cancer risk. Br J Cancer 1988;57:205–212.
11. Swanson CA, Potischman N, Wilbanks GD, Twiggs LB, Mortel R, Berman ML, Barrett RJ, Baumgartner RN, Brinton LA: Relation of endometrial cancer risk to past and contemporary body size and body fat distribution. Cancer Epidemiol Biomarkers Prev 1993;2:321–327.
12. Rosner B, Willett WC, Spiegelman D. Correction of logistic regression relative risk estimates and confidence intervals for systematic within-person measurement error. Stat Med 1989;8:1051–1069.
13. Voigt LF, Weiss NS. Epidemiology of endometrial cancer. Cancer Treat Res 1989;49:1–21.
14. Troisi R, Potischman N, Hoover RN, Siiteri P, Brinton LA and the Endometrial Cancer Collaborative Group. Insulin and endometrial cancer. Am J Epidemiol 1997;146:476–482.
15. Reaven GM. Role of insulin resistance in human disease (syndrome X): an expanded definition. Annu Rev Med 1993;234:461–469.
16. Hankinson SE, Manson JE, Spiegelman D, Willett WC, Longcope C, Speizer FE. Reproducibility of plasma hormone levels in postmenopausal women over a 2–3 year period. Cancer Epidemiol Biomarkers Prev 1995;4:649–654.
17. Wacholder S. When measurement errors correlate with truth: surprising effect of nondifferential misclassification. Epidemology 1995;6:157–161.

## Appendix: Calculation of Bias in the Estimate of the BMI Effect from Measurement Error in $\log_e$(Estrone)

Let $X_1$ be the true natural logarithm of estrone, $X_2$ be true BMI, and $X_3$ be the age for a given woman. We assume that the log odds of disease is $\mu + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3$ as in Eq 1 in the text. Suppose that BMI and age are measured without error but that a single measurement of log (estrone), $Z_1$, satisfies $Z_1 = X_1 + \epsilon$, where $\epsilon$ represents a measurement error, independent of $X_1$, $X_2$, $X_3$, with mean zero and variance $\sigma^2$. The variance, $\sigma^2$, represents both secular variation in log (estrone) within each woman and laboratory error. Letting $\Sigma = \{\sigma_{ij}\} = \{\text{cov}(X_i, X_j)\}$, assuming that $X_1$, $X_2$, $X_3$ and $\epsilon$ are jointly normal, and setting $Z_2 = X_2$, and $Z_3 = X_3$, we calculate the expected values

$$E\begin{pmatrix} X_1 \mid Z_1 \\ X_2 \mid Z_2 \\ X_3 \mid Z_3 \end{pmatrix} = \begin{pmatrix} Z_1 - \{\sigma^2\sigma^{11}(1 + \sigma^2\sigma^{11})\}\{\sigma^{11}(Z_1 - \mu_1) + \sigma^{12}(Z_2 - \mu_2) + \sigma^{13}(Z_3 - \mu_3)\} \\ Z_2 \\ Z_3 \end{pmatrix} \quad \text{(A1)}$$

where $\{\sigma^{ij}\}$ are elements in the inverse of the matrix $\Sigma$. In particular $\sigma^{11} = (\sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21})^{-1} \equiv \sigma_{11.2}^{-1}$, where $\Sigma_{12} = (\sigma_{12}, \sigma_{13})$ and $\Sigma_{22}$ is the covariance matrix of $(X_2, X_3)$. Replacing expected values of $X_1$, $X_2$, $X_3$ given $Z_1$, $Z_2$, $Z_3$ for $X_1$, $X_2$, $X_3$ in Eq 1, as in the regression calibration technique developed by Rosner, et al (12), we find that the coefficients $\beta_1^*$, $\beta_2^*$, and $\beta_3^*$ of $Z_1$, $Z_2$, and $Z_3$ are given by

$$\beta_1^* = \beta_1/(1 + \sigma^2\sigma^{11}) = \beta_1\{\sigma_{11.2}/(\sigma_{11.2} + \sigma^2)\}, \quad \text{(A2)}$$

$$\beta_2^* = \beta_2 - \beta_1\{\sigma^2\sigma_{11.2}/(\sigma^2 + \sigma_{11.2})\}\sigma^{12}, \quad \text{(A3)}$$

and

$$\beta_3^* = \beta_3 - \beta_1\{\sigma^2\sigma_{11.2}/(\sigma^2 + \sigma_{11.2})\}\sigma^{13}. \quad \text{(A4)}$$

These formulas extend in an obvious way for any number of independent variables $X_2, X_3, \ldots, X_k$ measured without error. From Eq A2, we find that $\beta_1^*$ is attenuated by a factor $\sigma_{11.2}/(\sigma_{11.2} + \sigma^2)$, because both $\sigma_{11.2}$ and $\sigma^2$ are non-negative. Of greater interest for our problem is that the coefficient, $\beta_2^*$, of $Z_2$ (the BMI measurement) is biased by $\beta_1\{\sigma^2\sigma_{11.2}/\sigma^2 + \sigma_{11.2})\}\sigma^{12}$. Because $\sigma^{12}$ is usually negative, $\beta_2^*$ will usually overestimate the true value $\beta_2$.

To estimate the bias in $\beta_2^*$, we calculated the sample covariance matrix of $Z_1$, $Z_2$, $Z$ from the sample of 209 postmenopausal women studied by Potischman *et al* (1) as $\hat{\sigma}^2 + \hat{\sigma}_{11} = 0.27435$, $\sigma_{12} = 1.14499$, $\sigma_{13} = -0.16992$, $\sigma_{22} = 37.38777$, $\sigma_{23} = -4.63604$, and $\sigma_{33} = 46.79062$. Other elements are given by symmetry. Under a nested random effects model, the variance of a $Z_1$ measurement is $\sigma_c^2 + \sigma_{11} + \sigma_b^2 + \sigma_r^2/r$, where $\sigma_c^2$ is the variance of a random batch effect (the women were analyzed in 14 batches), $\sigma_{11}$ is the variance of the true long term $\log_e$ (estrone) levels among these women, $\sigma_b^2$ is the variance from year to year variation in a given woman (see Ref 16), and $\sigma_r^2/r$ is the variance from laboratory replication error on a given sample with r replicates. An analysis variance test for $\sigma_c^2 = 0$ gave $F_{13,195} = 0.08$ with $p = 0.66$, and a moment estimator of $\sigma_c^2$ was $-0.0039$. We concluded that batch effects can be ignored in this analysis, and we set $\sigma_c^2 = 0$. As mentioned above, the variance of $Z_1$ was 0.27435, which we equate to $\hat{\sigma}_{11} + \hat{\sigma}_b^2 + \hat{\sigma}_r^2/r$. Hankinson *et al* (16) studied 80 postmenopausal women who each provided one blood sample in 1989–1990 and two subsequent blood samples spaced approximately at yearly intervals. Assuming no batch effect, we find that the intraclass correlation estimated by Hankinson *et al* is $\text{ICC} = \sigma_{11}/(\sigma_{11} + \sigma_b^2 + \sigma_r^2/r)$. Because Hankinson *et al* used the same laboratory

(Corning-Nichols Institute) as Potischman *et al*, it is reasonable to suppose $\sigma_c^2 = 0$ in the study by Hankinson *et al* and that $\sigma_\tau^2/r$ was the same in both studies. From the value ICC = 0.74 with 95% confidence interval (0.66, 0.83) given by Hankinson *et al*, we estimate $\hat{\sigma}_{11} = 0.74 \times 0.27435 = 0.20302$ and $\hat{\sigma}^2 \equiv \hat{\sigma}_b^2 + \hat{\sigma}_\tau^2/r = 0.27435 - 0.20302 = 0.07133$. From $\{\hat{\sigma}_{ij}\}$ we calculate $\hat{\sigma}_{11.2} = 0.16794$.

Estimates $\hat{\beta}_1^*$, $\hat{\beta}_2^*$, and $\hat{\beta}_3^*$ were obtained from Eq 1 with $Z_1$, $Z_2 = X_2$, and $Z_3 = X_3$ replacing $X_1$, $X_2$ and $X_3$, as 0.4374, 0.08092, and 0.02547, respectively. Using the previous estimates of $\sigma^2$, $\{\sigma^{ij}\}$, and $\sigma_{11.2}$ in Eq A2, A3, and A4, we invert these equations successively to obtain error-corrected estimates $\hat{\beta}_1 = 0.6232$, $\hat{\beta}_2 = 0.7524$ and $\hat{\beta}_3 = 0.02558$. Note that $\hat{\beta}_1$ exceeds $\hat{\beta}_1^*$ by 42%. More important for our purposes, $\hat{\beta}_2$ is less than $\hat{\beta}_2^*$ by an amount 0.0057, equivalent to a 7% decrease. As a sensitivity analysis, we assumed ICC = 0.66, which is the lower confidence limit of Hankinson *et al*, and obtained $\hat{\beta} = 0.07238$, which is 11% less than $\hat{\beta}_2^*$. Thus, measurement error in $Z_1$ does not explain the persistent effect of BMI after adjustment for $\log_e$ (estrone) in model (1).